

0570
1025

OIPE

#2

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/841,963

DATE: 11/13/2001

TIME: 15:37:56

Input Set : N:\Crf3\RULE60\09841963.txt

Output Set: N:\CRF3\11132001\I841963.raw

4 <110> APPLICANT: MUSC FOUNDATION FOR RESEARCH DEVELOPMENT
 6 <120> TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSIS AND TREATMENT
 7 OF CANCER BASED ON THE TRANSCRIPTION FACTOR ETS2
 9 <130> FILE REFERENCE: 9175-006-228
 11 <140> CURRENT APPLICATION NUMBER: 09/841,963
 12 <141> CURRENT FILING DATE: 2001-04-25
 14 <150> PRIOR APPLICATION NUMBER: 99/27805
 15 <151> PRIOR FILING DATE: 1999-11-23
 17 <160> NUMBER OF SEQ ID NOS: 34
 19 <170> SOFTWARE: FastSEQ for Windows Version 3.0
 21 <210> SEQ ID NO: 1
 22 <211> LENGTH: 2269
 23 <212> TYPE: DNA
 24 <213> ORGANISM: Homo sapiens
 26 <220> FEATURE:
 27 <221> NAME/KEY: CDS
 28 <222> LOCATION: (292)...(1698)
 30 <400> SEQUENCE: 1

| | | |
|----|---|---------|
| 31 | ccgtttcctc cctccctc cactcgccg tccctccttc ctctccctc ctccctcctc | 60 |
| 32 | ctcccgctcc tgaagagcgc gccgcgtggg ggaaggcccg gttacttctt ccagagactg | 120 |
| 33 | acgagtgcgg tgcgctcca gtcagagct cccggagccg cccggccagc gtcggcctc | 180 |
| 34 | cctgatcgct tctggccggc gccctcgccc tcgcccggcg cgcaccgagc agccgcgggc | 240 |
| 35 | gccgagcagc caccgtcccg accaagcgcc ggccctgccc gcagcggcag g atg aat | 297 |
| 36 | | Met Asn |
| 37 | | 1 |
| 39 | gat ttc gga atc aag aat atg gac cag gta gcc cct gtg gct aac agt | 345 |
| 40 | Asp Phe Gly Ile Lys Asn Met Asp Gln Val Ala Pro Val Ala Asn Ser | |
| 41 | 5 10 15 | |
| 43 | tac aga ggg aca ctc aag cgc cag cca gcc ttt gac acc ttt gat ggg | 393 |
| 44 | Tyr Arg Gly Thr Leu Lys Arg Gln Pro Ala Phe Asp Thr Phe Asp Gly | |
| 45 | 20 25 30 | |
| 47 | tcc ctg ttt gct gtt ttt cct tct cta aat gaa gag caa aca ctg caa | 441 |
| 48 | Ser Leu Phe Ala Val Phe Pro Ser Leu Asn Glu Glu Gln Thr Leu Gln | |
| 49 | 35 40 45 50 | |
| 51 | gaa gtg cca aca ggc ttg gat tcc att tct cat gac tcc gcc aac tgt | 489 |
| 52 | Glu Val Pro Thr Gly Leu Asp Ser Ile Ser His Asp Ser Ala Asn Cys | |
| 53 | 55 60 65 | |
| 55 | gaa ttg cct ttg tta acc ccg tgc agc aag gct gtg atg agt caa gcc | 537 |
| 56 | Glu Leu Pro Leu Leu Thr Pro Cys Ser Lys Ala Val Met Ser Gln Ala | |
| 57 | 70 75 80 | |
| 59 | tta aaa gct acc ttc agt ggc ttc aaa aag gaa cag cgg cgc ctg gcc | 585 |
| 60 | Leu Lys Ala Thr Phe Ser Gly Phe Lys Lys Glu Gln Arg Arg Leu Gly | |
| 61 | 85 90 95 | |
| 63 | att cca aag aac ccc tgg ctg tgg agt gag caa cag gta tgc cag tgg | 633 |
| 64 | Ile Pro Lys Asn Pro Trp Leu Trp Ser Glu Gln Gln Val Cys Gln Trp | |
| 65 | 100 105 110 | |
| 67 | ctt ctc tgg gcc acc aat gag ttc agt ctg gtg aac gtg aat ctg cag | 681 |

ENTERED

RAW SEQUENCE LISTING

DATE: 11/13/2001

PATENT APPLICATION: US/09/841,963

TIME: 15:37:56

Input Set : N:\Crf3\RULE60\09841963.txt

Output Set: N:\CRF3\11132001\I841963.raw

```

68 Leu Leu Trp Ala Thr Asn Glu Phe Ser Leu Val Asn Val Asn Leu Gln
69 115 120 125 130
71 agg ttc ggc atg aat ggc cag atg ctg tgt aac ctt ggc aag gaa cgc 729
72 Arg Phe Gly Met Asn Gly Gln Met Leu Cys Asn Leu Gly Lys Glu Arg
73 135 140 145
75 ttt ctg gag ctg gca cct gac ttt gtg ggt gac att ctc tgg gaa cat 777
76 Phe Leu Glu Leu Ala Pro Asp Phe Val Gly Asp Ile Leu Trp Glu His
77 150 155 160
79 ctg gag caa atg atc aaa gaa aac caa gaa aag aca gaa gat caa tat 825
80 Leu Glu Gln Met Ile Lys Glu Asn Gln Glu Lys Thr Glu Asp Gln Tyr
81 165 170 175
83 gaa gaa aat tca cac ctc acc tcc gtt cct cat tgg att aac agc aat 873
84 Glu Glu Asn Ser His Leu Thr Ser Val Pro His Trp Ile Asn Ser Asn
85 180 185 190
87 aca tta ggt ttt ggc aca gag cag gcg ccc tat gga atg cag aca cag 921
88 Thr Leu Gly Phe Gly Thr Glu Gln Ala Pro Tyr Gly Met Gln Thr Gln
89 195 200 205 210
91 aat tac ccc aaa ggc ggc ctc ctg gac agc atg tgt ccg gcc tcc aca 969
92 Asn Tyr Pro Lys Gly Gly Leu Leu Asp Ser Met Cys Pro Ala Ser Thr
93 215 220 225
95 ccc agc gta ctc agc tct gag cag gag ttt cag atg ttc ccc aag tct 1017
96 Pro Ser Val Leu Ser Ser Glu Gln Glu Phe Gln Met Phe Pro Lys Ser
97 230 235 240
99 cgg ctc agc tcc gtc agc gtc acc tac tgc tct gtc agt cag gac ttc 1065
100 Arg Leu Ser Ser Val Ser Val Thr Tyr Cys Ser Val Ser Gln Asp Phe
101 245 250 255
103 cca ggc agc aac ttg aat ttg ctc acc aac aat tct ggg act ccc aaa 1113
104 Pro Gly Ser Asn Leu Asn Leu Leu Thr Asn Asn Ser Gly Thr Pro Lys
105 260 265 270
107 gac cac gac tcc cct gag aac ggt gcg gac agc ttc gag agc tca gac 1161
108 Asp His Asp Ser Pro Glu Asn Gly Ala Asp Ser Phe Glu Ser Ser Asp
109 275 280 285 290
111 tcc ctc ctc cag tcc tgg aac agc cag tcg tcc ttg ctg gat gtg caa 1209
112 Ser Leu Leu Gln Ser Trp Asn Ser Gln Ser Ser Leu Leu Asp Val Gln
113 295 300 305
115 cgg gtt cct tcc ttc gag agc ttc gaa gat gac tgc agc cag tct ctc 1257
116 Arg Val Pro Ser Phe Glu Ser Phe Glu Asp Asp Cys Ser Gln Ser Leu
117 310 315 320
119 tgc ctc aat aag cca acc atg tct ttc aag gat tac atc caa gag agg 1305
120 Cys Leu Asn Lys Pro Thr Met Ser Phe Lys Asp Tyr Ile Gln Glu Arg
121 325 330 335
123 agt gac cca gtg gag caa ggc aaa cca gtt ata cct gca gct gtg ctg 1353
124 Ser Asp Pro Val Glu Gln Gly Lys Pro Val Ile Pro Ala Ala Val Leu
125 340 345 350
127 gcc ggc ttc aca gga agt gga cct att cag ctg tgg cag ttt ctc ctg 1401
128 Ala Gly Phe Thr Gly Ser Gly Pro Ile Gln Leu Trp Gln Phe Leu Leu
129 355 360 365 370
131 gag ctg cta tca gac aaa tcc tgc cag tca ttc atc agc tgg act gga 1449
132 Glu Leu Leu Ser Asp Lys Ser Cys Gln Ser Phe Ile Ser Trp Thr Gly

```

RAW SEQUENCE LISTING

DATE: 11/13/2001

PATENT APPLICATION: US/09/841,963

TIME: 15:37:56

Input Set : N:\Crf3\RULE60\09841963.txt

Output Set: N:\CRF3\11132001\I841963.raw

```

133          375          380          385
135 gac gga tgg gag ttt aag ctc gcc gac ccc gat gag gtg gcc cgc cgg      1497
136 Asp Gly Trp Glu Phe Lys Leu Ala Asp Pro Asp Glu Val Ala Arg Arg
137          390          395          400
139 tgg gga aag agg aaa aat aag ccc aag atg aac tac gag aag ctg agc      1545
140 Trp Gly Lys Arg Lys Asn Lys Pro Lys Met Asn Tyr Glu Lys Leu Ser
141          405          410          415
143 cgg ggc tta cgc tac tat tac gac aag aac atc atc cac aag acg tcg      1593
144 Arg Gly Leu Arg Tyr Tyr Tyr Asp Lys Asn Ile Ile His Lys Thr Ser
145          420          425          430
147 ggg aag cgc tac gtg tac cgc ttc gtg tgc gac ctc cag aac ttg ctg      1641
148 Gly Lys Arg Tyr Val Tyr Arg Phe Val Cys Asp Leu Gln Asn Leu Leu
149          435          440          445          450
151 ggg ttc acg ccc gag gaa ctg cac gcc atc ctg ggc gtc cag ccc gac      1689
152 Gly Phe Thr Pro Glu Glu Leu His Ala Ile Leu Gly Val Gln Pro Asp
153          455          460          465
155 acg gag gac tgaggtcgcc gggaccaccc tgagccggcc ccaggtcgt      1738
156 Thr Glu Asp
159 ggactgagtg ggaagcccat cctgaccagc tgcctccgag gaccaggaag aggcaggatt      1798
160 gaaaatgtcc agaaaagtgg ccaagaagca gtggccttat tgcaccccaa accacgcctc      1858
161 ttgaccaggc tgcctccctt gtggcagcaa cggcacagct aattctactc acagtgcctt      1918
162 taagtgaataa tggctgagaa agaggcaccg ggaagccgtc ctggcgccctg gcagtcctgt      1978
163 ggacgggatg gttctggctg tttgagattc tcaaaggagc gagcatgtcg tggacacaca      2038
164 cagactatatt ttagattttt ttttgccttt tgcaaccagg aacagcaaat gcaaaaactc      2098
165 tttgagaggg taggaggggtg ggaaggaaac aaccatgtca ttccagaagt tagttgtat      2158
166 atattataat aatcttataa ttgttctcag aatcccttaa cagttgtatt taacagaaat      2218
167 tgtatattgt aattttaaata aattatataa ctgtatttga aataagaatt c      2269
169 <210> SEQ ID NO: 2
170 <211> LENGTH: 469
171 <212> TYPE: PRT
172 <213> ORGANISM: Homo sapiens
174 <400> SEQUENCE: 2
175 Met Asn Asp Phe Gly Ile Lys Asn Met Asp Gln Val Ala Pro Val Ala
176 1 5 10 15
177 Asn Ser Tyr Arg Gly Thr Leu Lys Arg Gln Pro Ala Phe Asp Thr Phe
178 20 25 30
179 Asp Gly Ser Leu Phe Ala Val Phe Pro Ser Leu Asn Glu Glu Gln Thr
180 35 40 45
181 Leu Gln Glu Val Pro Thr Gly Leu Asp Ser Ile Ser His Asp Ser Ala
182 50 55 60
183 Asn Cys Glu Leu Pro Leu Leu Thr Pro Cys Ser Lys Ala Val Met Ser
184 65 70 75 80
185 Gln Ala Leu Lys Ala Thr Phe Ser Gly Phe Lys Lys Glu Gln Arg Arg
186 85 90 95
187 Leu Gly Ile Pro Lys Asn Pro Trp Leu Trp Ser Glu Gln Gln Val Cys
188 100 105 110
189 Gln Trp Leu Leu Trp Ala Thr Asn Glu Phe Ser Leu Val Asn Val Asn
190 115 120 125
191 Leu Gln Arg Phe Gly Met Asn Gly Gln Met Leu Cys Asn Leu Gly Lys

```

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/841,963

DATE: 11/13/2001
 TIME: 15:37:56

Input Set : N:\Crf3\RULE60\09841963.txt
 Output Set: N:\CRF3\11132001\I841963.raw

```

192      130      135      140
193 Glu Arg Phe Leu Glu Leu Ala Pro Asp Phe Val Gly Asp Ile Leu Trp
194 145      150      155      160
195 Glu His Leu Glu Gln Met Ile Lys Glu Asn Gln Glu Lys Thr Glu Asp
196      165      170      175
197 Gln Tyr Glu Glu Asn Ser His Leu Thr Ser Val Pro His Trp Ile Asn
198      180      185      190
199 Ser Asn Thr Leu Gly Phe Gly Thr Glu Gln Ala Pro Tyr Gly Met Gln
200      195      200      205
201 Thr Gln Asn Tyr Pro Lys Gly Gly Leu Leu Asp Ser Met Cys Pro Ala
202      210      215      220
203 Ser Thr Pro Ser Val Leu Ser Ser Glu Gln Glu Phe Gln Met Phe Pro
204 225      230      235      240
205 Lys Ser Arg Leu Ser Ser Val Ser Val Thr Tyr Cys Ser Val Ser Gln
206      245      250      255
207 Asp Phe Pro Gly Ser Asn Leu Asn Leu Leu Thr Asn Asn Ser Gly Thr
208      260      265      270
209 Pro Lys Asp His Asp Ser Pro Glu Asn Gly Ala Asp Ser Phe Glu Ser
210      275      280      285
211 Ser Asp Ser Leu Leu Gln Ser Trp Asn Ser Gln Ser Ser Leu Leu Asp
212      290      295      300
213 Val Gln Arg Val Pro Ser Phe Glu Ser Phe Glu Asp Asp Cys Ser Gln
214 305      310      315      320
215 Ser Leu Cys Leu Asn Lys Pro Thr Met Ser Phe Lys Asp Tyr Ile Gln
216      325      330      335
217 Glu Arg Ser Asp Pro Val Glu Gln Gly Lys Pro Val Ile Pro Ala Ala
218      340      345      350
219 Val Leu Ala Gly Phe Thr Gly Ser Gly Pro Ile Gln Leu Trp Gln Phe
220      355      360      365
221 Leu Leu Glu Leu Leu Ser Asp Lys Ser Cys Gln Ser Phe Ile Ser Trp
222      370      375      380
223 Thr Gly Asp Gly Trp Glu Phe Lys Leu Ala Asp Pro Asp Glu Val Ala
224 385      390      395      400
225 Arg Arg Trp Gly Lys Arg Lys Asn Lys Pro Lys Met Asn Tyr Glu Lys
226      405      410      415
227 Leu Ser Arg Gly Leu Arg Tyr Tyr Tyr Asp Lys Asn Ile Ile His Lys
228      420      425      430
229 Thr Ser Gly Lys Arg Tyr Val Tyr Arg Phe Val Cys Asp Leu Gln Asn
230      435      440      445
231 Leu Leu Gly Phe Thr Pro Glu Glu Leu His Ala Ile Leu Gly Val Gln
232      450      455      460
233 Pro Asp Thr Glu Asp
234 465
236 <210> SEQ ID NO: 3
237 <211> LENGTH: 84
238 <212> TYPE: PRT
239 <213> ORGANISM: Homo sapiens
241 <400> SEQUENCE: 3
242 Ser Gly Pro Ile Gln Leu Trp Gln Phe Leu Leu Glu Leu Leu Thr Asp

```

RAW SEQUENCE LISTING

DATE: 11/13/2001

PATENT APPLICATION: US/09/841,963

TIME: 15:37:56

Input Set : N:\Crf3\RULE60\09841963.txt

Output Set: N:\CRF3\11132001\I841963.raw

```

243      1              5              10              15
244 Lys Ser Cys Gln Ser Phe Ile Ser Trp Thr Gly Asp Gly Trp Glu Phe
245              20              25              30
246 Lys Leu Ser Asp Pro Asp Glu Val Ala Arg Arg Trp Gly Lys Arg Lys
247              35              40              45
248 Asn Lys Pro Lys Met Asn Tyr Glu Lys Leu Ser Arg Gly Leu Arg Tyr
249              50              55              60
250 Tyr Tyr Asp Lys Asn Ile Ile His Lys Thr Ala Gly Lys Arg Tyr Val
251 65              70              75              80
252 Tyr Arg Phe Val
255 <210> SEQ ID NO: 4
256 <211> LENGTH: 84
257 <212> TYPE: PRT
258 <213> ORGANISM: Mus musculus
260 <400> SEQUENCE: 4
261 Ser Gly Pro Ile Gln Leu Trp Gln Phe Leu Leu Glu Leu Leu Thr Asp
262 1              5              10              15
263 Lys Ser Cys Gln Ser Phe Ile Ser Trp Thr Gly Asp Gly Trp Glu Phe
264              20              25              30
265 Lys Leu Ser Asp Pro Asp Glu Val Ala Arg Arg Trp Gly Lys Arg Lys
266              35              40              45
267 Asn Lys Pro Lys Met Asn Tyr Glu Lys Leu Ser Arg Gly Leu Arg Tyr
268              50              55              60
269 Tyr Tyr Asp Lys Asn Ile Ile His Lys Thr Ala Gly Lys Arg Tyr Val
270 65              70              75              80
271 Tyr Arg Phe Val
274 <210> SEQ ID NO: 5
275 <211> LENGTH: 84
276 <212> TYPE: PRT
277 <213> ORGANISM: Gallus gallus
279 <400> SEQUENCE: 5
280 Ser Gly Pro Ile Gln Leu Trp Gln Phe Leu Leu Glu Leu Leu Thr Asp
281 1              5              10              15
282 Lys Ser Cys Gln Ser Phe Ile Ser Trp Thr Gly Asp Gly Trp Glu Phe
283              20              25              30
284 Lys Leu Ser Asp Pro Asp Glu Val Ala Arg Arg Trp Gly Lys Arg Lys
285              35              40              45
286 Asn Lys Pro Lys Met Asn Tyr Glu Lys Leu Ser Arg Gly Leu Arg Tyr
287              50              55              60
288 Tyr Tyr Asp Lys Asn Ile Ile His Lys Thr Ala Gly Lys Arg Tyr Val
289 65              70              75              80
290 Tyr Arg Phe Val
293 <210> SEQ ID NO: 6
294 <211> LENGTH: 84
295 <212> TYPE: PRT
296 <213> ORGANISM: Xenopus laevis
298 <400> SEQUENCE: 6
299 Ser Gly Pro Ile Gln Leu Trp Gln Phe Leu Leu Glu Leu Leu Thr Asp
300 1              5              10              15

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/841,963

DATE: 11/13/2001

TIME: 15:37:57

Input Set : N:\Crf3\RULE60\09841963.txt
Output Set: N:\CRF3\11132001\I841963.raw